Modeling dominated choice alternatives using the constrained multinomial logit

F. Martinez, University of Chile, CL; E. Cascetta, F. Pagliara, University of Naples Federico II, IT, M. Bierlaire, EPFL Lausanne, CH, K. W. Axhausen, ETH, Zurich, CH

Speaker: Michel Bierlaire

transp-or.epfl.ch

Transport and Mobility Laboratory
Ecole Polytechnique Fédérale de Lausanne, Switzerland
Outline

- Motivation
- Concept of cutoffs (Constrained logit model)
- Concept of dominance
- Using dominance in the Constrained Logit Model
- Preliminary results
- Perspectives
Motivation

- Discrete choice models.
- Concept of utility based on trade-offs.
- Attributes threshold generally not accounted for.
- Dominated alternatives may not even be considered in the choice set.
- How do we model that?
Motivation

- Manski (1977): individual-based choice-set based on deterministic constraints
- Swait and Ben-Akiva (1987): random constraints
- Swait (2001), Martinez et al. (2008): Attribute cutoffs
- Cascetta and Papola (2005), Cascetta et al. (2007): implicit perception, dominance values

Idea: combine cutoffs and dominance
Cutoffs

Optimization problem of rational consumer $n$:

$$\max_{\delta_{ni}} \sum_{i \in C} \delta_{ni} U_{in}(X_i)$$

subject to

$$\sum_{i \in C} \delta_{ni} = 1, \quad \delta_{ni} \in \{0, 1\}, \forall i \in C$$

But attributes are meaningful only within some bounds

$$\ell_{nk} \leq X_{ik} \leq u_{nk} \forall i \in C, \forall k$$
Cutoffs

Idea: relax the constraint in a probabilistic way

Example: constraint $\ell \leq X$

\[
\begin{align*}
V_{\text{not considered}} &= \ell + \varepsilon_1 \\
V_{\text{considered}} &= X + \varepsilon_2
\end{align*}
\]

\[
P(\text{considered}) = \frac{e^{\rho X}}{e^{\rho X} + e^{\rho \ell}} = \frac{1}{1 + e^{\rho (\ell - X)}}
\]

Example: constraint $X \leq u$

\[
P(\text{considered}) = \frac{e^{-\rho X}}{e^{-\rho X} + e^{-\rho u}} = \frac{1}{1 + e^{\rho (X - u)}}
\]
Cutoffs

Example: $2 \leq X$
Cutoffs

Example: $X \leq 4$
Cutoffs

Constraint $\ell \leq X \leq u$

$$P(\text{considered}) = \frac{1}{1 + e^{\rho(\ell - X)}} \frac{1}{1 + e^{\rho(X - u)}}$$

We denote this quantity by $\phi_n(X)$
Cutoffs

Example: $2 \leq X \leq 4$
Cutoffs

The utility function now becomes

\[ V_i = \sum_k \beta_k X_{ik} + \sum_{k^*} \frac{1}{\rho} \ln \phi_n(X_{ik^*}) \]

where \( k^* \) ranges only on constrained attributes. Note that

\[ \ln \phi(X) = - \ln(1 + e^{\rho(\ell - X)}) - \ln(1 + e^{\rho(X - u)}) \]

\[ = - \ln(1 + e^{\rho \ell} e^{-\rho X}) - \ln(1 + e^{\rho X} e^{-\rho u}) \]

Can be estimated, although it is difficult
Dominance

- Destination choice (origin $o$)
- Dominance variables: reflect the spatial position and hierarchies of alternatives
- Dominance rules:
  - Weak dominance: Alternative $d$ dominates alternative $d^*$ if
    1. $A_d > A_{d^*}$ (attractivity attribute)
    2. $c_{od} < c_{od^*}$ (generalized transportation cost)
  - Strong dominance: $d$ strongly dominates $d^*$ if it weakly dominates it and is along the path to reach $d^*$ from $o$
Dominance

\[ WP_1 = WP_2 = WP_3 = WP_4 \]
\[ c_{OD1} = c_{OD2} = c_{OD3} < c_{OD4} \]

Dₑ, D₂, D₃ dominate WEAKLY D₄

D₂ STRONGLY dominates D₄

area of possible zones STRONGLY dominating D₄
**Dominance**

Examples of dominance variables for destination \( d \). Consider 3 conditions:

(a) \( d^* \) has average price lower than \( d \)
(b) \( \text{dist}(o, d^*) < \text{dist}(o, d) \)
(c) Strong rule: \( \text{dist}(o, d^*) + \text{dist}(d^*, d) < \text{dist}(o, d) \)

**Strong global dominance variable**  nbr of \( d^* \) verifying (a), (b) and (c).

**Weak global dominance variable**  nbr of \( d^* \) verifying (a) and (b)

**Weak spatial dominance variable**  nbr of \( d^* \) verifying (b)

**Strong spatial dominance variable**  nbr of \( d^* \) verifying (b) and (c).
Dominance

Dominance variables are introduced directly in the utility function of an MNL model (Cascetta and Papola, 2005):

$$U_d = \sum_k \beta_k X_{dk} + \sum_j \gamma_j Y_{dj}$$
Dominance within CML

Idea: alternatives with a high dominance variable are not considered

Constraint:

\[ Y_{dj} \leq u \]

Problem: what is a reasonable threshold \( u \)?

Let’s use the cutoffs:

\[
\ln \phi(Y_{dj}) = -\ln(1 + e^{\rho Y_{dj}} e^{-\rho u}) = -\ln(1 + \bar{u} e^{\rho Y_{dj}})
\]

We try to estimate \( \bar{u} \)
Case study: canton Zürich
## Residential location choice

Model specification:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Price}_d$</td>
<td>average land price of zone $d$</td>
</tr>
<tr>
<td>$\text{LnStock}_d$</td>
<td>log of the housing stock in zone $d$</td>
</tr>
<tr>
<td>$\text{Logsum}_{od}^{\text{LM}}$</td>
<td>logsum of the mode choice model for work purpose (low-medium income)</td>
</tr>
<tr>
<td>$\text{Logsum}_{od}^{H}$</td>
<td>logsum of the mode choice model for work purpose (high income)</td>
</tr>
<tr>
<td>$\text{LnWorkPlacesServ}_d$</td>
<td>log of the workplaces in services (retail, leisure, services, incl. education and health) in $d$. Measure of quality of services.</td>
</tr>
</tbody>
</table>
Number of observations = 657

\[ \mathcal{L}(0) = -3419.032 \]

\[ \mathcal{L}(\hat{\beta}) = -53.971 \]

\[ -2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 6730.123 \]

\[ \rho^2 = 0.984 \]

\[ \bar{\rho}^2 = 0.983 \]

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Logsum(^H)(_{od})</td>
<td>15.3</td>
<td>2.85</td>
</tr>
<tr>
<td>2</td>
<td>Logsum(^LM)(_{od})</td>
<td>16.6</td>
<td>2.97</td>
</tr>
<tr>
<td>3</td>
<td>Price(_d)</td>
<td>-0.00160</td>
<td>0.000221</td>
</tr>
<tr>
<td>4</td>
<td>LnStock(_d)</td>
<td>1.12</td>
<td>0.102</td>
</tr>
<tr>
<td>5</td>
<td>LnWorkPlacesServ(_d)</td>
<td>0.187</td>
<td>0.180</td>
</tr>
</tbody>
</table>
MNL

- Very high $\rho^2$: 0.98
- Correct signs
- Significant parameters, except the level of services

Next model:
- Include the strong spatial dominance variable (based only on distance, not on price)
- Simple linear specification

$$V_d = \cdots + \beta \text{dom}_d$$
Linear dominance

Number of observations = 657

\[ L(0) = -3419.032 \]

\[ L(\hat{\beta}) = -47.055 \]

\[ -2[L(0) - L(\hat{\beta})] = 6743.955 \]

\[ \rho^2 = 0.986 \]

\[ \bar{\rho}^2 = 0.984 \]

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \text{dom}_d )</td>
<td>-0.0859</td>
<td>0.0120</td>
</tr>
<tr>
<td>2</td>
<td>( \text{Logsum}_{od}^H )</td>
<td>16.1</td>
<td>2.62</td>
</tr>
<tr>
<td>3</td>
<td>( \text{Logsum}_{od}^{LM} )</td>
<td>17.1</td>
<td>2.76</td>
</tr>
<tr>
<td>4</td>
<td>( \text{Price}_d )</td>
<td>-0.00245</td>
<td>0.000313</td>
</tr>
<tr>
<td>5</td>
<td>( \text{LnStock}_d )</td>
<td>1.20</td>
<td>0.133</td>
</tr>
<tr>
<td>6</td>
<td>( \text{LnWorkPlacesServ}_d )</td>
<td>-0.172</td>
<td>0.198</td>
</tr>
</tbody>
</table>
Linear dominance

- Significantly better fit: $-2(-53.971 - 47.055) = 202.052$
- Correct signs
- Significant parameters, except the level of services

Next model: cutoff

$$V_d = \cdots - \ln(1 + \bar{u} \exp(\rho \text{dom}_d))$$

$$= \cdots - \ln(1 + 1000 \exp(\rho \text{dom}_d))$$

Notes:

- the estimation of $\bar{u}$ failed; its value continuously increased
- in the final model, the value $\bar{u} = 1000$ was used.
Cutoff

Number of observations = 657

\[ \mathcal{L}(0) = -3419.032 \]

\[ \mathcal{L}(\hat{\beta}) = -47.057 \]

\[ -2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 6743.952 \]

\[ \rho^2 = 0.986 \]

\[ \bar{\rho}^2 = 0.984 \]

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt. std. error</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Logsum\textsuperscript{H}\textsubscript{od}</td>
<td>16.1</td>
<td>2.62</td>
<td>6.16</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>Logsum\textsuperscript{LM}\textsubscript{od}</td>
<td>17.1</td>
<td>2.76</td>
<td>6.20</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>Price\textsubscript{d}</td>
<td>-0.00245</td>
<td>0.000313</td>
<td>-7.82</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>LnStock\textsubscript{d}</td>
<td>1.20</td>
<td>0.133</td>
<td>9.01</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>LnWorkPlacesServ\textsubscript{d}</td>
<td>-0.172</td>
<td>0.198</td>
<td>-0.87</td>
<td>0.39</td>
</tr>
<tr>
<td>6</td>
<td>\rho</td>
<td>0.0859</td>
<td>0.0120</td>
<td>7.17</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Cutoff

- Same improvement than the linear specification
- Actually, the model is almost linear, due to the high value of $\bar{u}$
- Question: can we accept a linear specification?
- We test it using a Box-Cox transform.

$$V_d = \cdots + \beta \frac{\text{dom}_d^\lambda - 1}{\lambda}$$
Box-Cox test

Number of observations = 657

\[
\mathcal{L}(0) = -3419.032 \\
\mathcal{L}(\hat{\beta}) = -43.120 \\
-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 6751.826 \\
\rho^2 = 0.987 \\
\bar{\rho}^2 = 0.985
\]

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dom&lt;sub&gt;d&lt;/sub&gt;</td>
<td>-0.579</td>
<td>-10.74 0.00</td>
</tr>
<tr>
<td>2</td>
<td>Logsum&lt;sub&gt;od&lt;/sub&gt;H</td>
<td>16.9</td>
<td>6.36 0.00</td>
</tr>
<tr>
<td>3</td>
<td>Logsum&lt;sub&gt;od&lt;/sub&gt;LM</td>
<td>18.0</td>
<td>6.72 0.00</td>
</tr>
<tr>
<td>4</td>
<td>Price&lt;sub&gt;d&lt;/sub&gt;</td>
<td>-0.00292</td>
<td>-9.00 0.00</td>
</tr>
<tr>
<td>5</td>
<td>LnStock&lt;sub&gt;d&lt;/sub&gt;</td>
<td>1.42</td>
<td>8.10 0.00</td>
</tr>
<tr>
<td>6</td>
<td>LnWorkPlacesServ&lt;sub&gt;d&lt;/sub&gt;</td>
<td>-0.328</td>
<td>-1.28 0.20</td>
</tr>
<tr>
<td>7</td>
<td>(\lambda)</td>
<td>0.434</td>
<td>11.19 0.00</td>
</tr>
</tbody>
</table>
Box-Cox test

- $\lambda$ is significantly different from 1.0 ($t$-test = 14.6)
- $\lambda$ is significantly different from 0.0 ($t$-test = 11.2)
- The linear specification is rejected
Conclusions

- Main idea: combination of two concepts: cutoffs and dominance
- First estimation results produces large values for the variance of the cutoff, so that it is basically equivalent to the linear model
- But... the linear specification is clearly rejected by a formal test.
- Next steps:
  - Consider new dominance rules, more consistent with the use of cutoffs
  - Investigate other data sets